Architectural Considerations in Delivering a Balanced Linux Cluster

Douglas Doerfler

Scalable Systems Integration Dept.

Sandia National Labs

Balance of the Entire System

- What is Cplant?
- Balance with respect to:
  - User Needs
  - Research & Development Needs
  - System Administration & Operations
  - Management Tools
  - Integration & Test

What is Cplant?

Cplant™ is a concept

- Computational capacity at a low cost
- MPPs using commodity components

Cplant™ is multiple efforts

- Scalable System Software
- Message Passing Protocols
- Integration and Testing
- Scalable System Management Tools

Cplant™ is a software package

- Downloadable Open Source
- Unlimited Scale Inc.: Commercial License

Cplant™ is a Philosophy

- Models ASCI Red Architecture
- MPP "Look and Feel"
- Preserve Application Code Base

ASCI Red

CPlant™

- "Scale Down" Design
- Leverage System Support Resources

NM clusters

- Antarctica: (FY00/01)
  Four independent heads plus a swinging center section, 2,528 total nodes
- Siberia (reapplied): 592 nodes, (FY99)
- Alaska: 272 nodes (FY98)
- Barrow (retired): 96 nodes (FY98)

•CA clusters

- Delmar: 384 nodes (FY00/01)
- Zenia: 32 nodes (FY00)
- Carmel: 128 nodes (FY99)
- Asilomar-SON: 64 nodes (FY97)
- Asilomar-SRN: 64 nodes (FY97)

Cplant is Multi-Discipline

Collaborations

- DOE SciDAC Program
- University of New Mexico
- Ohio State
- Syracuse
- University of Virginia
- University of Texas
- Mississippi State
- Northwestern
- Cal Tech/JPL

Supporting System Software Efforts

- Unlimited Scale Inc.
- MPI Software Technology
- Etnus Inc.
- Kuck & Associates/Pallas

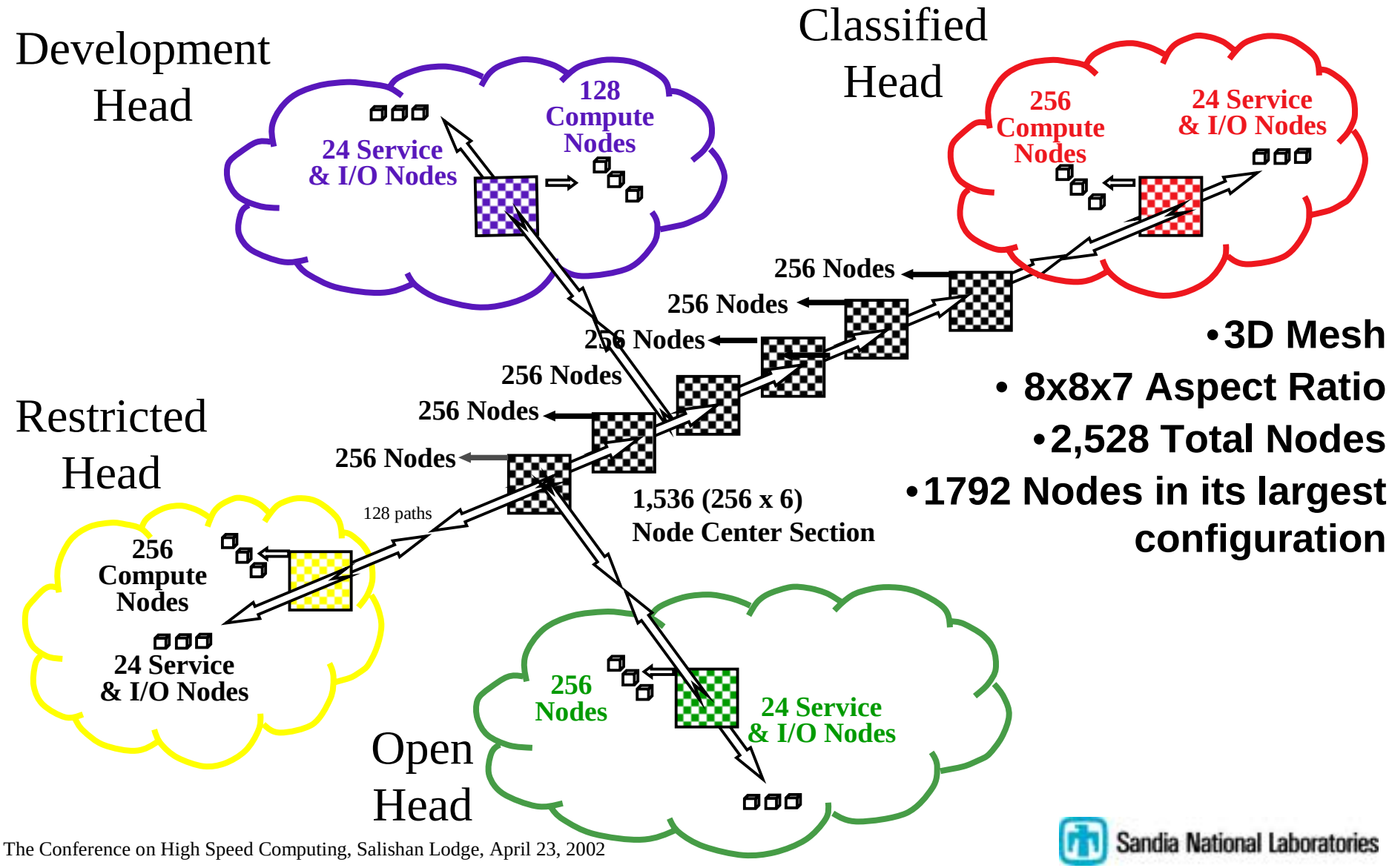
Applications

- ALEGRA
- SALINAS
- PRONTO
- CTH
- LAMMPS
- PARADYN
- LADERA
- QUEST
- …

Multi-Program

- DOE DP
- DOE Non-DP
- WFO
- Research Foundation
- LDRD

Antarctica



- 3D Mesh
- 8x8x7 Aspect Ratio
- 2,528 Total Nodes
- 1792 Nodes in its largest configuration

Sandia National Laboratories

Antarctica Status

- **West Head (SON):** Deployed March '01 (80 nodes), Nov '01 (256 nodes)
  - 96 Compaq DS10L (Slate):
    466Mhz EV6 Alpha CPU, 1GByte Memory, Myrinet 2000
  - 160 Compaq XP1000:
    500Mhz EV6 Alpha CPU, 1GByte Memory, Myrinet 2000
  - 24 Compaq XP1000 Service and I/O nodes
- **Center Section (Swings Between Heads):** Deployed May '01 (1024 nodes), Sept '01 (1536 nodes)
  - 1536 Compaq DS10L:
    466Mhz EV6 Alpha CPU, 256MByte Memory, Myrinet 1280
- **Ross Head (SRN):** Deployed Nov '01
  - 256 Compaq DS10L (Slate) Compute Nodes:
    600Mhz EV67 Alpha CPU, 1GByte Memory, Myrinet 2000
  - 24 Compaq XP1000 Service and I/O nodes
- **Ronne Head (SCN):** Deployed March '02
  - 256 Compaq XP1000 Compute Nodes:
    500Mhz EV6 Alpha CPU, 1GByte Memory, Myrinet 1280
  - 24 Compaq XP1000 Service and I/O nodes

User Needs Vs. Development Needs

- An eclectic mix of users
  - Three types of networks at Sandia
  - An Institutional Resource: Post Docs to Plimpton
- An eclectic mix of developers
  - Research Platform
  - Development Platform
  - Integration & Test
  - System Administrators

Production Challenges

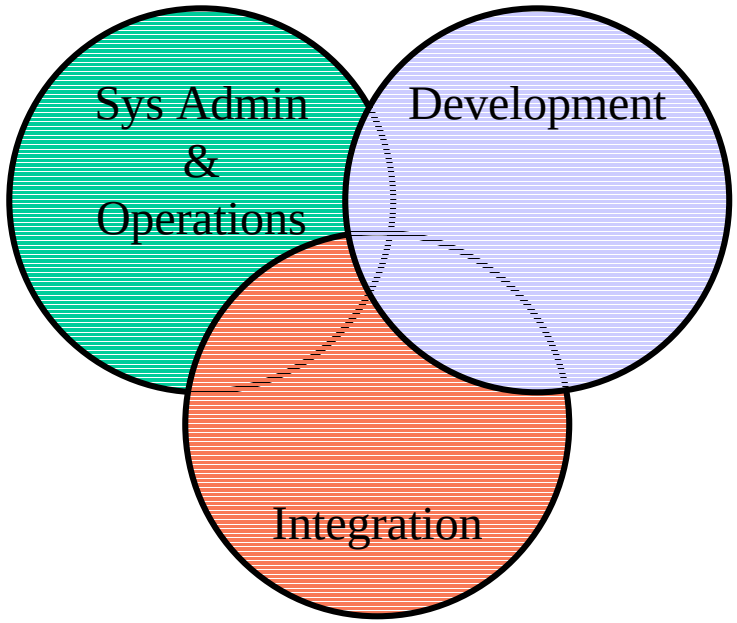
- Customers
  - Cplant is being positioned as an institutional resource
    - Wide variety of user expertise, codes, …
    - High profile exposure
  - Early adopters turned off by the rough edges. As we get better how do we gain back their confidence?
  - Users expect support on the level of commercial machines, especially ASCI Red and White
  - Bad news travels at the speed of light.
  - Good news doesn't travel at all!
  - We're working on our PR and marketing departments!

Development Challenges

- **Availability: Users vs. Those $%*@ Developers**
  - **How do we balance the needs of the users & increase the quality of the product?**
    - **Debug scalability problems**
    - **Continue our fundamental research efforts**
    - **Scheduled Maintenance**
- **Presently**
  - **2 days every 2 weeks for System Time**
  - **User dedicated time upon request**

Balanced Level of Effort

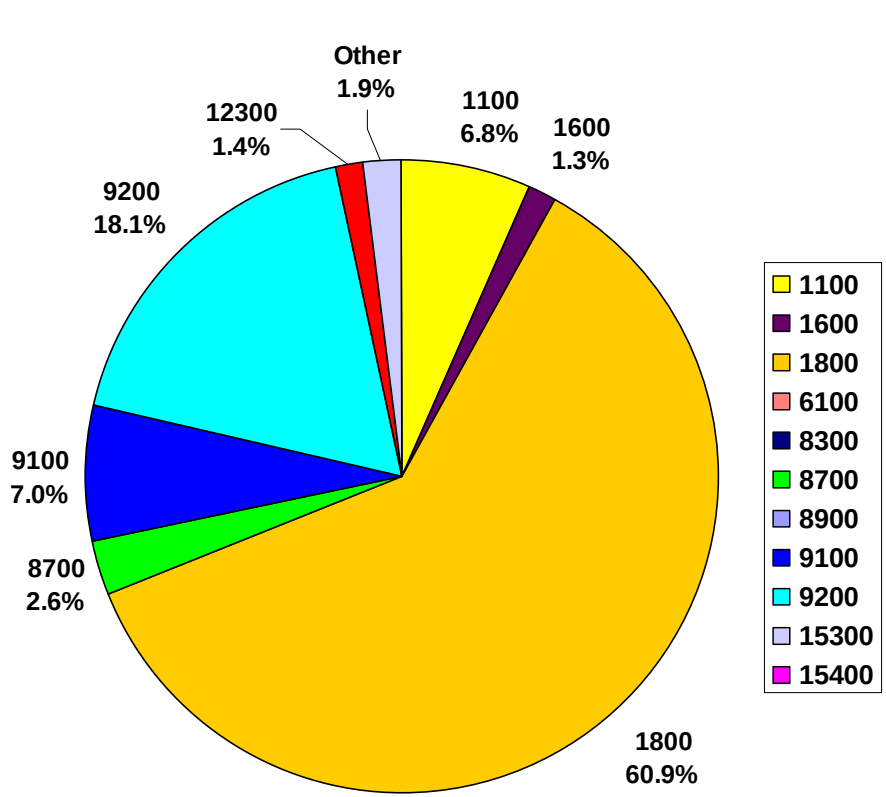
- System Administration and Operations
  - Day to day management
  - cplant-help, email help line
  - Scheduled maintenance
  - Problem resolution
  - Computational Scientists
  - Notifications
- Integration and Development
  - Software support
  - Hardware problem resolution
  - SW and HW diagnostic development



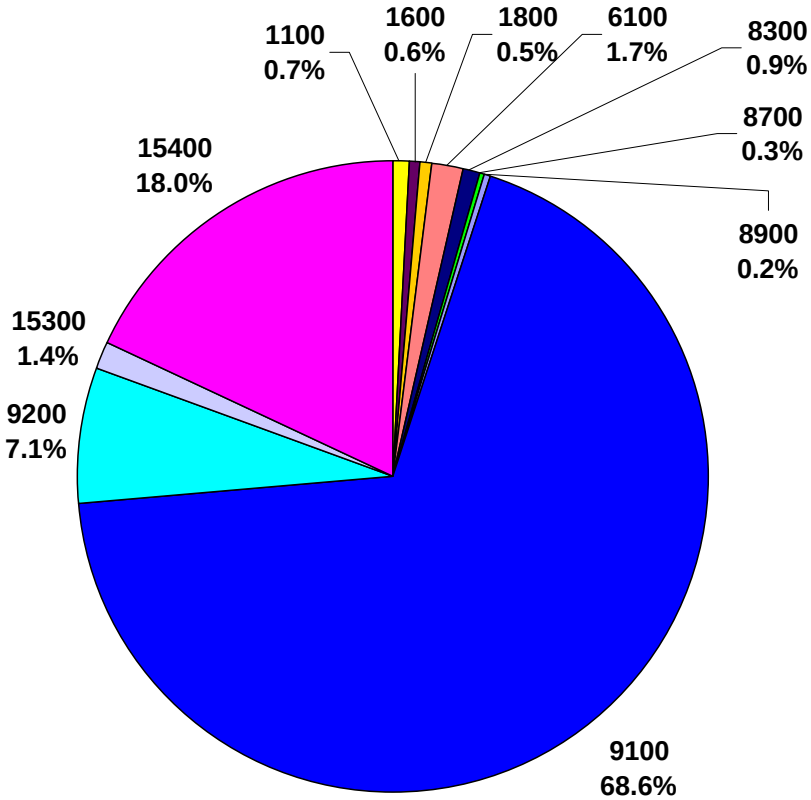
Ongoing Support Level of Effort

- Department 9338 (7 FTE ?)
  - Operations, Security, Computational Scientists
- Department 9223 (7 FTE)
  - Run Time Software Dev. and Support, Computational Scientist
- Department 9224 (4 FTE)
  - Integration & Test, Cluster Management Tool Dev. and Support
- Compaq
  - Professional Services (3 FTE)
  - Hardware diagnosis and repair (HW Contract)
  - Compiler & Operating System Support (SW License)
- High Performance Technologies Inc. (1.5 FTE)
  - Integration, Diagnostics, Cluster Management Tools

Cplant & ASCI Red Usage by Org (Dec '01 thru Feb '02)



Cplant



ASCI Red

The Cluster Management Challenge

- Operational Staff w/o in-depth machine knowledge
- Heterogenous mix of hardware
  - Compute, Service, I/O, Power, Terminals, Networks, …
- Classification switching
- Scaling Issues:
  - 1,000's of nodes
  - Boot/Reboot time
  - Easily target a specific device
- Runtime independence
- Portability & Extensibility

Some requirements

- Support Disk-less AND Disk-full nodes
- Support multiple software environments at the node level
- Support switching between classified and unclassified networks
- Support a hierarchical administrative (diagnostic) network
- Separate management tools and runtime environment.
- Manage cluster as a single system
  - Doesn't mean single system image
- Do not require kernel modifications
- Do not effect performance of compute nodes
- Be usable by cluster no-experts
- Boot in less than half an hour
- Any more?

What we don't know

- What type of devices we will use in the future
  - What capabilities
- How we will connect these devices together in the future
- How these clusters will be used in the future
- Can we design something now that we won't have to throw away?

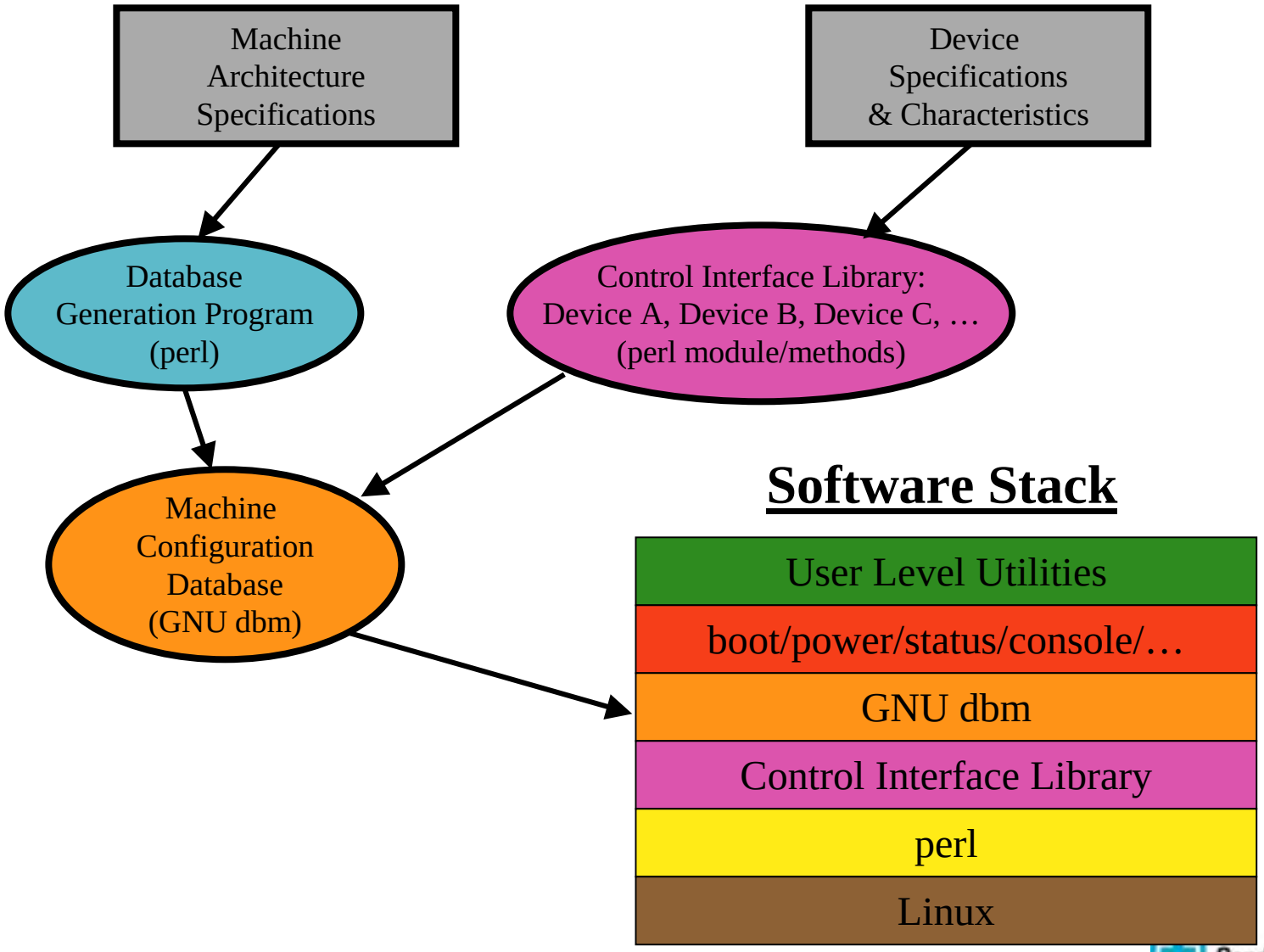
What should be provided?

- power and boot control
- General health of device (again not just nodes)
  - Nodes
    - Up or down?
    - Interface status (all interfaces)
    - Memory errors
    - Temp
    - Fans
  - Other devices
    - Any info that is important to Reliability Availability Serviceability
- Anything that any device needs to perform their function
  - Firmware, kernels, root-file-system….
- Single point of control
  - Could be many single points?
  - That single point could move?

Sandia's Approach

- Don't be focused ONLY on Sandia's needs
- Don't make assumptions like:
  - What devices or how they are connected
- First effort was diskless
  - We needed it for switching
  - Many benefits anyway
- Use what we can from open source
  - Return what we create

Cplant Cluster Management Tools



Management Tools Summary

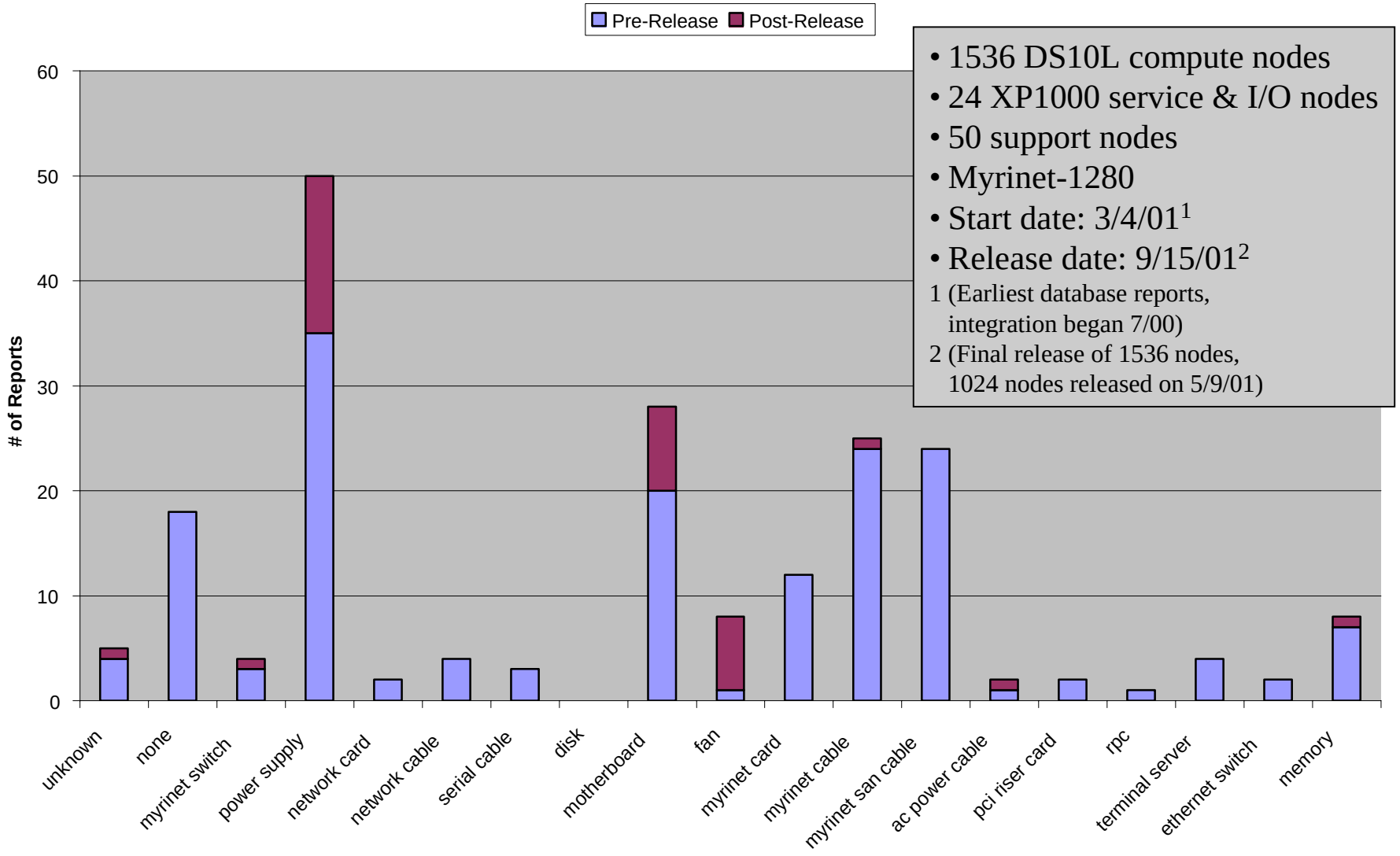
- Many clusters installed at Sandia and HPTi
  - All different in either device make up, topology or both
  - Largest 1,861 nodes, single hard-drive
- We believe the toolset is flexible enough to target a wider audience than Cplant
- Open source model similar to our runtime software distributions
  - But independent of the runtime

Integration & Test

(How do you verify that the 1,000's of components in the system operate per specification and are connected correctly?)

- **Integration and Test have become key elements of our cluster effort**
- **Cluster Integration Toolkit (CIToolkit)**
  - **A collaborative effort between Sandia & HPTi**
  - **Functional tests for *every* HW component**
  - **Connection verification**
  - **Stress tests, primarily memory and network**
  - **Topology verification**
- **System Software Installation (OS and Runtime)**
  - **Stress tests**
  - **Algorithm testing**
  - **Friendly Users**

Center Section Hardware Reports

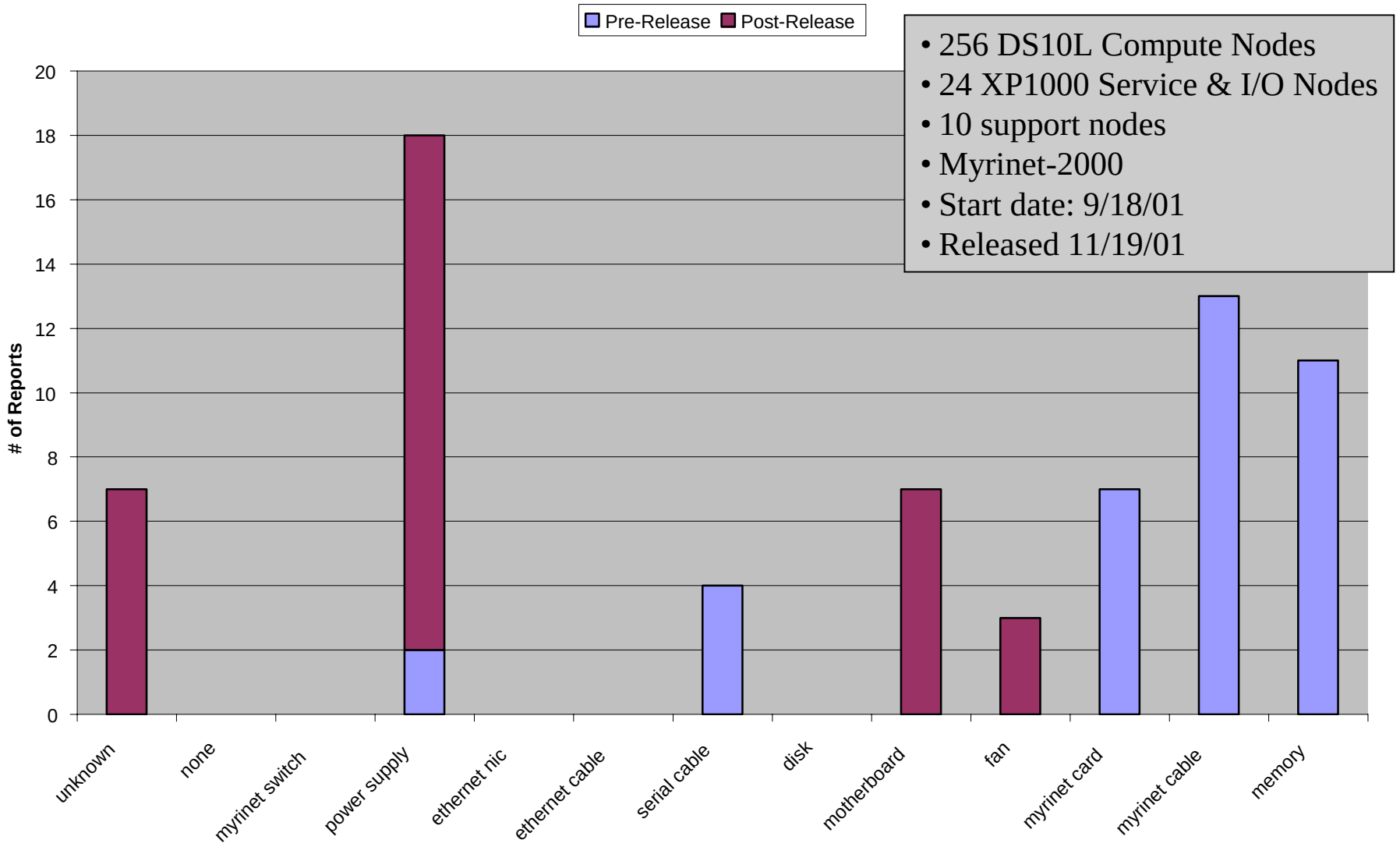


- 1536 DS10L compute nodes
- 24 XP1000 service & I/O nodes
- 50 support nodes
- Myrinet-1280
- Start date: 3/4/01[1]
- Release date: 9/15/01[2]

1 (Earliest database reports, integration began 7/00)

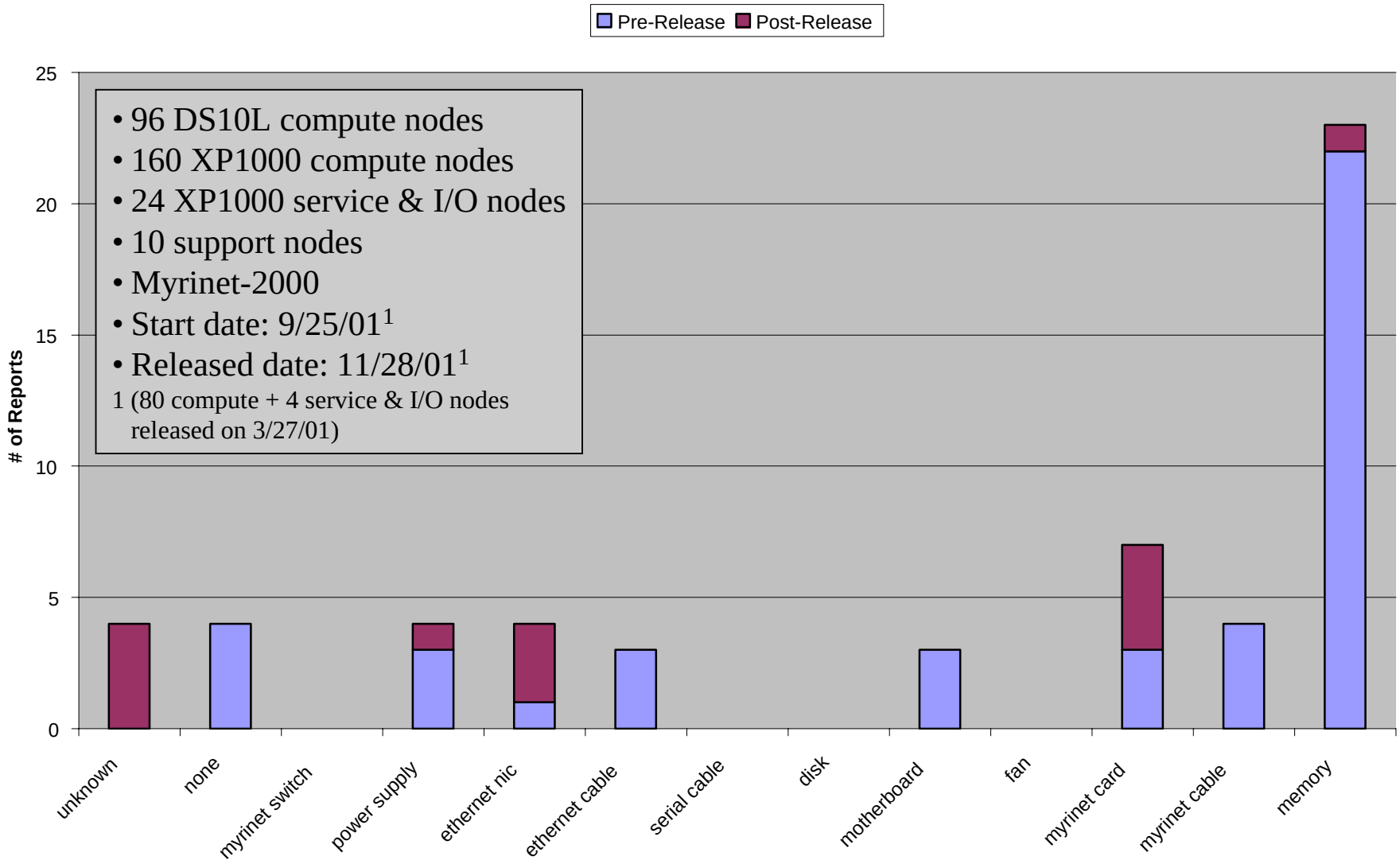
2 (Final release of 1536 nodes, 1024 nodes released on 5/9/01)

Ross Hardware Reports



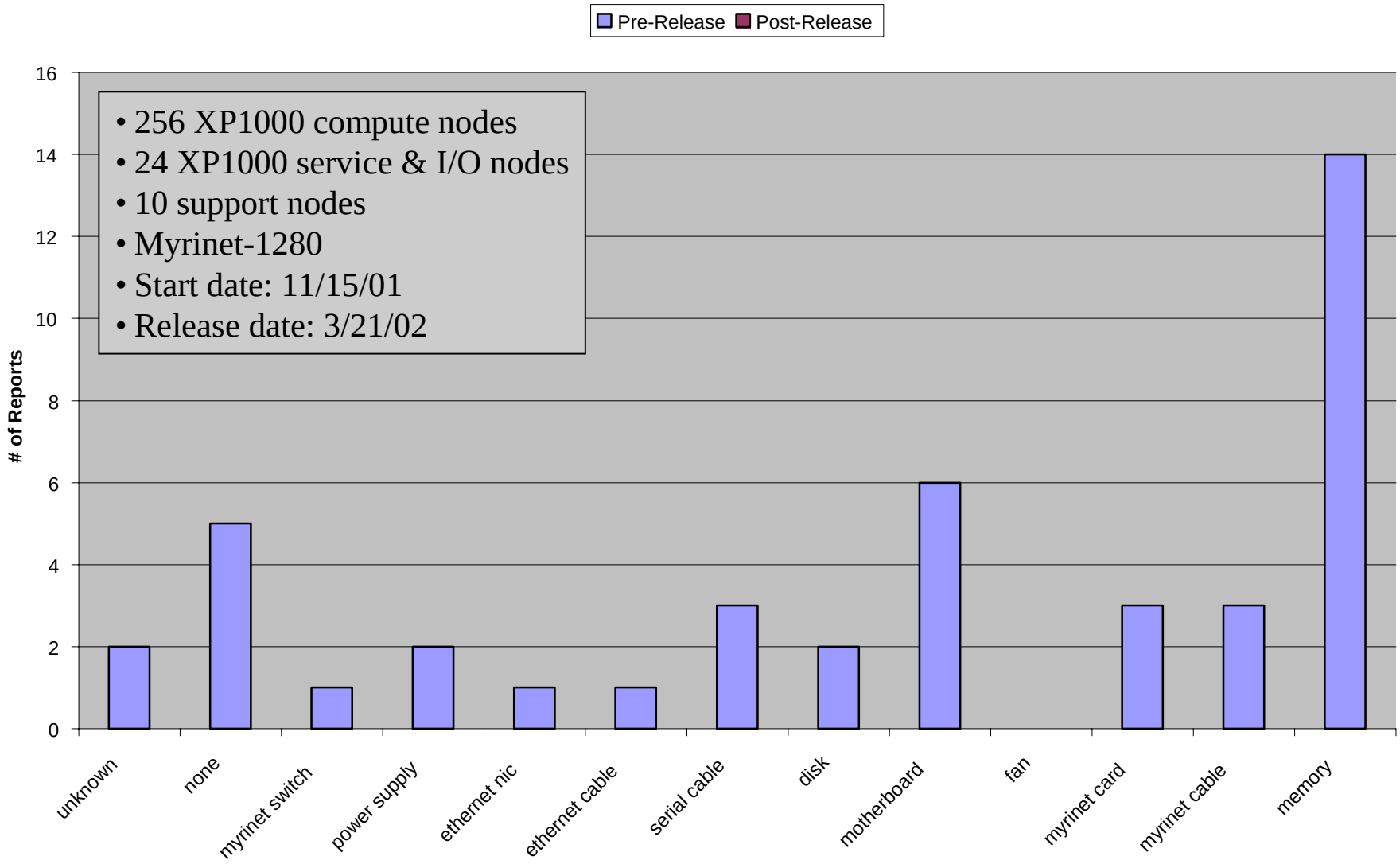
- 256 DS10L Compute Nodes
- 24 XP1000 Service & I/O Nodes
- 10 support nodes
- Myrinet-2000
- Start date: 9/18/01
- Released 11/19/01

West Hardware Reports



Sandia National Laboratories

Ronne Hardware Reports



Sandia National Laboratories

More Info

Web site
http://www.cs.sandia.gov/cplant

Open Source
http://www.cs.sandia.gov/cplant/download.html

Recent papers
http://www.cs.sandia.gov/cplant/papers